APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/684,508	10/15/2003	Livia Polanyi	CQ10224	6736

23493 7590 12/09/2008		EXAMINER
SUGHRUE MION, PLLC		COLUCCI, MICHAEL C
2100 Pennsylvania Avenue, N.W.		
Washington, DC 20037	ART UNIT	PAPER NUMBER
	2626	

NOTIFICATION DATE	DELIVERY MODE
12/09/2008	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

USPTO@sughrue.com
USPatDocketing@sughrue.com

	Application No.	Applicant(s)
Office Action Summary	10/684,508	POLANYI ET AL.
	Examiner	Art Unit
	MICHAEL C. COLUCCI	2626

-- *The MAILING DATE of this communication appears on the cover sheet with the correspondence address* --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE <u>3</u> MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
  Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) ☐ Responsive to communication(s) filed on _____.

2a) ☐ This action is **FINAL**.   2b) ☒ This action is non-final.

3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) ☒ Claim(s) <u>1-23 and 34-47</u> is/are pending in the application.

     4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5) ☒ Claim(s) <u>2-9,12-21,34,35 and 37</u> is/are allowed.

6) ☒ Claim(s) <u>1,10,11,22,23 and 36</u> is/are rejected.

7) ☐ Claim(s) _____ is/are objected to.

8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) ☐ The specification is objected to by the Examiner.

10) ☒ The drawing(s) filed on <u>15 October 2003</u> is/are: a)☒ accepted or b)☐ objected to by the Examiner.

     Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

     Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12) ☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

     a)☐ All   b)☐ Some * c)☒ None of:

       1.☐ Certified copies of the priority documents have been received.

       2.☐ Certified copies of the priority documents have been received in Application No. _____.

       3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

     * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☐ Information Disclosure Statement(s) (PTO/SB/08)
Paper No(s)/Mail Date _____.

4) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____.

5) ☐ Notice of Informal Patent Application

6) ☐ Other: _____.

APPLICATION NO./ CONTROL NO.	FILING DATE	FIRST NAMED INVENTOR / PATENT IN REEXAMINATION	ATTORNEY DOCKET NO.
10684508	10/15/2003	POLANYI ET AL.	CQ10224

EXAMINER	
RICHEMOND DORVIL	

ART UNIT	PAPER
2626	20081204

SUGHRUE MION, PLLC
401 Castro Street, Ste 220
Mountain View, CA 94041-2007

DATE MAILED:

Please find below and/or attached an Office communication concerning this application or proceeding.

Commissioner for Patents

Richemond Dorvil
SPE
Art Unit: 2626

PTO-90C (Rev.04-03)

DETAILED ACTION

NOTE: This action has **not** been made final due merely to recent court decisions regarding claims not being tied to a statutory class or transforming subject matter. See 35 U.S.C. 101 rejection below.

Response to Arguments

1.   Applicant's arguments filed 10/09/2008 have been fully considered but they are not persuasive.

Argument 1 (page 18):

- "However, the portion of the claim element reciting "based on at least one non-structural measure of relevance" is still not accounted for in Corston."

- "As such, Corston does not teach "relevance scores for the discourse constituents based on at least one non-structural measure of relevance" of claim 1"

- "As such, Corston fails to teach or suggest elements of "determining relevance scores ... percolating the relevance scores" of claim 1 for which it is cited."

Response to argument 1:

Examiner takes the position that Marcu in view of Corston in fact teaches the limitations of the present invention, particularly claim 1. Examiner would like to point out the parallel teachings of document retrieval in both the present invention and cited references (present invention [0013]), wherein keywords from

retrieved documents are analyzed, and a score is calculated to differentiate

information in a document. Like Corston, the present invention utilizes a

repository of documents to perform discourse analysis on. Thus, Corston

teaches user input query in such techniques is typically presented as either an

explicit user generated query, or an implicit query, such as when <u>a user requests</u>

<u>documents which are similar to a set of existing documents</u>. Typical <u>information</u>

<u>retrieval systems search documents in the larger data store at either a single</u>

<u>word level, or at a term level</u>. Each of the documents is assigned <u>a relevancy (or</u>

<u>similarity) score, and the information retrieval system presents a certain subset of</u>

<u>the documents searched to the user, typically that subset which has a relevancy</u>

<u>score which exceeds a given threshold</u> (Corston Col. 1 line 59 – Col. 2 line 33).

Regarding non-structured, Corston overcomes well known limitations of previous

inventions, wherein Corston teaches statistical search engines stems from their

assumption that words are independent variables, i.e., words in any textual

passage occur independently of each other. Independence in this context means

that a conditional probability of any one word appearing in a document given the

presence of another word therein is always zero, i.e., <u>a document simply</u>

<u>contains an unstructured collection of words or simply put a "bag of words"</u>. As

one can readily appreciate, this assumption, with respect to any language, is

grossly erroneous. <u>English, like other languages, has a rich and complex</u>

<u>syntactic and lexicon-semantic structure with words whose meanings vary</u>, often

widely, <u>based on the specific linguistic context in which they are used</u>, with the context determining in any one instance a given meaning of a word and what word(s) can subsequently appear. Hence, w<u>ords that appear in a textual passage are simply not independent of each other, rather they are highly inter-dependent</u> (Corston Col. 1 line 59 – Col. 2 line 33). Thus, Corston analyzes unstructured text with a more sophisticated form, wherein context is used.

Regard both structured and unstructured text, wherein unstructured text can be input text hand typed by a user in an email, word processor, and even audio, Marcu teaches the present application relates to computational linguistics and more particularly to techniques for parsing a text to determine its <u>underlying rhetorical, or discourse, structure, and to techniques for summarizing, or compressing, text</u> ([0003]).

Further, Marcu teaches discourse parsing systems and techniques as described herein. The systems and techniques described here result in <u>a discourse parsing system that uses a set of learned decision rules to automatically determine the underlying discourse structure of any unrestricted text</u>. As a result, <u>the discourse parsing system can be used, among other ways, for constructing discourse trees whose leaves are sentences (or units that can be identified at high levels of performance)</u>. Moreover, the time, expense, and inconsistencies associated with manually built discourse tree derivation rules are reduced dramatically ([0019]).

Furthermore, Marcu teaches the formation of unstructured text into structured text, wherein Marcu teaches a decision-based rhetorical parsing system (equivalently, a discourse parsing system) <u>automatically derives the discourse structure of unrestricted texts and incrementally builds corresponding discourse trees based on a set of learned decision rules</u>. The discourse parsing system uses a shift-reduce rhetorical parsing algorithm that <u>learns to construct rhetorical structures of texts from a corpus of discourse-parse action sequences</u>. The rhetorical parsing algorithm implements robust lexical, syntactic and semantic knowledge sources ([0067]).

Claim Rejections - 35 USC § 101

2.   35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

Claims 1-10 are rejected under 35 U.S.C. 101 because:

Claims 1-10 fail to clearly recite a statutory process to which it is tied. Claims 1-10 recite purely mental steps and would not qualify as a statutory process. In order to qualify as a statutory process, the method claim should positively recite the other statutory class to which it is tied (i.e. apparatus, device, product, etc.). For example, the method steps of claims 1, 5, 6 ,and 8 appear to recite mental steps such as "determining a hybrid text summary" and do not identify an apparatus that performs the

recited method steps, such as "a client device" as described in the specification (present

invention [0128]).

Claim Rejections - 35 USC § 103

3.   The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set
> forth in section 102 of this title, if the differences between the subject matter sought to be patented and
> the prior art are such that the subject matter as a whole would have been obvious at the time the
> invention was made to a person having ordinary skill in the art to which said subject matter pertains.
> Patentability shall not be negatived by the manner in which the invention was made.

4.   Claims 1, 11, 22, 23, and 36 are rejected under 35 U.S.C. 103(a) as being

unpatentable over Marcu et al U.S. PGPUB 20020046018 (herein after Marcu) in view

of Corston et al. US 6901399 B1 (hereinafter Corston).

Re claims 1, 11, 22, 23, and 36, Marcu teaches a method of determining a hybrid

text summary comprising the steps of:

determining discourse constituents for a text ([0010]);

determining a structural representation of discourse for the text ([0003]);

determining relevance scores for the discourse constituents based on at least

one non- structural measure of relevance ([0003]);

determining a hybrid text summary ([0003])

However, Marcu fails to teach determining relevance scores and percolating the

relevance scores

Corston teaches that it is well know to use a relevancy score, wherein a user input query in such techniques is typically presented as either an explicit user generated query, or an implicit query, such as when a user requests documents which are similar to a set of existing documents. Typical information retrieval systems search documents in the larger data store at either a single word level, or at a term level. Each of the documents is assigned a relevancy (or similarity) score, and the information retrieval system presents a certain subset of the documents searched to the user, typically that subset which has a relevancy score which exceeds a given threshold (Corston Col.. 1 line 59 – Col.. 2 line 2).

Corston teaches statistical search engines stems from their assumption that words are independent variables, i.e., words in any textual passage occur independently of each other. Independence in this context means that a conditional probability of any one word appearing in a document given the presence of another word therein is always zero, i.e., a document simply contains an unstructured collection of words or simply put a "bag of words". As one can readily appreciate, this assumption, with respect to any language, is grossly erroneous. English, like other languages, has a rich and complex syntactic and lexicon-semantic structure with words whose meanings vary, often widely, based on the specific linguistic context in which they are used, with the context determining in any one instance a given meaning of a word and what word(s) can subsequently appear. Hence, words that appear in a textual passage are simply not independent of each other, rather they are highly inter-dependent (Corston Col. 1 line 59 – Col. 2 line 33).

based on the structural representation of discourse

Corston teaches that even if the document does not have a title which contains the keyword, sentences in the document can be analyzed to determine the meta structure of the document. For example, the subjects of sentences, particularly subjects of sentences whose main verb is "be", tend to be the theme or topic of that sentence. Precision can be increased, even for keyword queries, by preferentially matching the keyword queries against documents containing sentences about that keyword (Corston Col.. 33 line 65 – Col.. 34 line 9).

Corston teaches that a meta structure of the document is indicative of a general subject matter of the document, wherein obtaining the set of abstract logical forms indicative of a meta structure of the document comprises: obtaining the set of abstract logical forms based on formatting information corresponding to the document. Additional, obtaining the set of abstract logical forms indicative of a meta structure of the document comprises: obtaining the set of abstract logical forms based on topics of sentences in the document. Further, Corston teaches that obtaining the set of abstract logical forms indicative of a meta structure of the document comprises: obtaining the set of abstract logical forms based on subjects of sentences in the document (Corston Col.. 45 lines 37-55).

Therefore, it would have been obvious to one of ordinary skill in the art at the time of the invention determining a hybrid text summary based on constituents with relevance scores having a threshold score, where perCol.ation is performed based on a

relevance score. Percolating or passing relevant scores would allow for efficient

clustering, where only a specific range of structured data that abides by relevant criteria

will remain.

5.   **Claim 10 is rejected under 35 U.S.C. 103(a) as being unpatentable over**

Marcu et al U.S. PGPUB 20020046018 (herein after Marcu) in view of Corston et al.

US 6901399 B1 (hereinafter Corston) and further in view of Nakao US 6205456 B1

(hereinafter Nakao).

Re claim 10, Marcu fails to teach the method of claim 1, wherein percolation of

relevance scores comprises the steps of:

Corston teaches percolating relevance score of important discourse constituents

exceeding an importance threshold value (Corston Col.. 1 line 59 – Col.. 2 line 2).

Corston teaches that it is well know to use a relevancy score, wherein a user

input query in such techniques is typically presented as either an explicit user generated

query, or an implicit query, such as when a user requests documents which are similar

to a set of existing documents. Typical information retrieval systems search documents

in the larger data store at either a single word level, or at a term level. Each of the

documents is assigned a relevancy (or similarity) score, and the information retrieval

system presents a certain subset of the documents searched to the user, typically that

subset which has a relevancy score which exceeds a given threshold (Corston Col.. 1

line 59 – Col.. 2 line 2).

Therefore, it would have been obvious to one of ordinary skill in the art at the

time of the invention determining a hybrid text based on constituents with percolating

relevance scores having a threshold score, where threshold comparison is performed

based on a relevance score. Percolating relevance scores relevant to a threshold score

would allow for efficient clustering, where only a specific range of data that abides by a

relevance probability will remain.

However, Marcu in view of Corston fails to teach determining important discourse

constituents;

determining unresolved anaphors of the important discourse constituents (Nakao

Col. 4 line 12-29);

determining potential resolving discourse constituents containing potential

antecedent referent of the unresolved anaphors (Nakao Col. 4 line 12-29);

through the reduced span of potential resolving discourse constituents (Nakao

Col. 4 line 12-29);

determining a modified span of the important discourse constituents based on

the relevance score (Nakao Col. 4 line 12-29).

Nakao teaches for an anaphoric expression, its antecedent is searched for and

the anaphoric expression is replaced with the antecedent or a portion containing the

antecedent is included in a summary so that the summary can be easily understood.

The antecedent of the anaphoric expression can be identified by a method referred to

as a centering method. This method makes a list of centers that comprises probable

elements (centers) of a sentence to be antecedents of anaphoric expressions in the

subsequent sentences. The elements probability to be an antecedent is calculated

mainly by its syntactic role in a sentence, such as subject, direct object, etc. Then, the

method resolves an anaphoric expression by selecting the most probable element from

the list with the restriction of agreement of number, gender, etc.

NOTE: For purposes of prior art, a reduced span is construed to be both

functionally equivalent and equally effective as a resolved set with the most probable

elements present.

Therefore, it would have been obvious to one of ordinary skill in the art at the

time of the invention determining portions of text containing anaphors, where an

antecedent referent can be substituted for the anaphor, and removing the antecedent

referent relevant to a relevancy threshold. Using an anaphor to replace an antecedent

referent would allow for a preservation of memory, where more than one item can refer

to a single item (i.e. anaphoric relationship). Additionally, using syntactic analysis with

antecedent and anaphoric substitutions would allow for a minimized set of data that

would the most probable with respect to the discourse/text by reduction of a set of

candidate data. Additionally, percolating scores of various elements relevant to

probability and antecedent basis would allow for a further reduced set that is minimized

for both anaphor and original data from the discourse.

Allowable Subject Matter

6.   Claims 2-9, 12-21, 34, 35, and 37 are allowed.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Michael C. Colucci whose telephone number is (571)-270-1847. The examiner can normally be reached on 9:30 am - 6:00 pm, Monday-Friday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on (571)-272-7602. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Michael C Colucci/
Examiner, Art Unit 2626
Patent Examiner
AU 2626

(571)-270-1847

Michael.Colucci@uspto.gov

/Richemond Dorvil/

Supervisory Patent Examiner, Art Unit 2626